

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-222688

(43) 公開日 平成10年(1998) 8月21日

(51) Int.Cl.<sup>6</sup>

識別記号

F I

G 0 6 T 11/60

G 0 6 F 15/62

3 2 5 A

7/00

15/70

3 3 0 Q

H 0 4 N 1/40

H 0 4 N 1/40

F

1/409

1 0 1 C

審査請求 未請求 請求項の数3 F D (全 13 頁)

(21) 出願番号 特願平10-23781

(22) 出願日 平成10年(1998) 1月21日

(31) 優先権主張番号 7 8 6 5 3 8

(32) 優先日 1997年 1月21日

(33) 優先権主張国 米国 (U S)

(71) 出願人 590000798

ゼロックス コーポレイション

XEROX CORPORATION

アメリカ合衆国 ニューヨーク州 14644

ロチェスター ゼロックス スクエア

(番地なし)

(72) 発明者 デニス エル. ペナブル

アメリカ合衆国 14505 ニューヨーク州

マリオン ドーメディー ヒル ロード

4353

(74) 代理人 弁理士 中島 淳 (外1名)

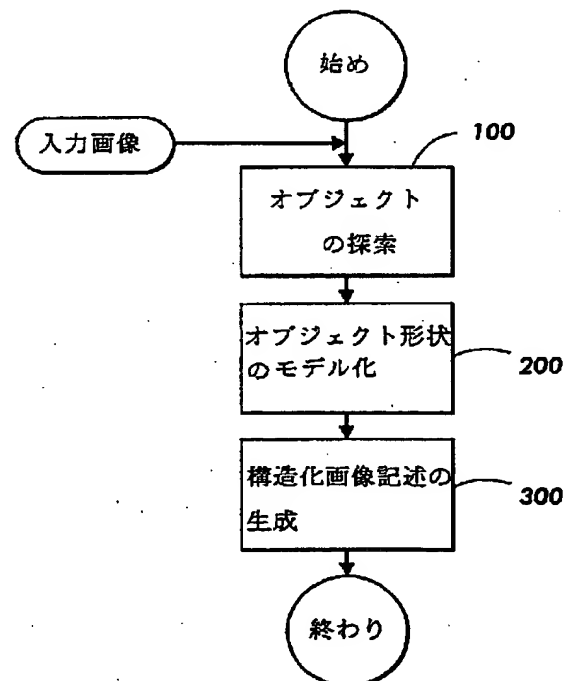
最終頁に続く

(54) 【発明の名称】 画像処理方法

(57) 【要約】

【課題】 画像中のオブジェクトを特徴づける画像処理方法を提供する。

【解決手段】 本発明に係る画像処理方法によれば、スキャナーのプラテン上に並べられた複数の異なる写真をスキャンすることによって生成された入力画像中の各写真の位置、形状、方向角を自動的に認識する。入力画像の処理中にコンピュータ22によって実行されるプロセスは、3つの一般的工程を備える。まず、ステップ100では、画像中のオブジェクトが探索して、そのオブジェクト境界を識別する。ステップ200では、オブジェクト形状をモデル化する。ステップ300では、その画像とオブジェクトを表現する構造化画像を生成する。その構造化画像は好適には画像データ自体を表現するデータだけでなく、各オブジェクトの位置、形状、方向やそれらの組み合わせを表現するデータを含む。



## 【特許請求の範囲】

【請求項 1】 デジタル入力画像を処理してその中の複数のオブジェクトを特徴づける方法であって、前記入力画像中で少なくとも 2 つのオブジェクトを識別する工程と、前記オブジェクトの各々の境界を表現する形状をモデル化する工程と、前記オブジェクトを特徴づける記述を生成する工程とを備える画像処理方法。

【請求項 2】 前記少なくとも 2 つのオブジェクトを識別する工程は、少なくとも 2 つのオブジェクトを囲む背景領域を識別する工程と、形態素フィルタリング処理を使って画像中のノイズ性エッジを平滑化する工程と、連続した前景領域を探索する工程とを備える請求項 1 に記載の画像処理方法。

【請求項 3】 前記背景領域を識別する工程は、2 値マスクを生成する工程であって、前記マスク中の各位置は前記入力画像のピクセルを表わす工程と、背景色を決定する工程と、適応シードフィル処理を使って、もし前記色が前記背景色に実質的に等しければ前記マスク中の各 2 進位置を第 1 の状態に設定し、そうでないならば、それを第 2 の状態に設定する工程と、必要ならば前記背景色を調整する工程と、少なくとも全オブジェクトのエッジが識別されるまで前記工程を繰り返す工程とを備える請求項 2 に記載の画像処理方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、一般的な画像処理技術に関し、特に、画像入力装置のプラテン上に置かれた複数のイメージ体を自動的に分割（セグメント化）して特徴を抽出する技術に関する。

## 【0002】

【従来の技術】一回のスキャンで取り込まれた画像が、好ましくはスキャナーのプラテン上に接触しないで並べて置かれた複数の異なる写真からなる場合、各写真の位置、形、回転角を自動的に確認することが望まれる。その性能によって生産性を上げることができる。多数の画像をスキャンする時間を短くし、並びの補正を自動化することによって生産性を向上させることができる。

## 【0003】

【発明が解決しようとする課題】本発明は、スキャンされた画像中の個々の領域、即ち、セグメントを自動的に決定するために開発されたソフトウェアに基づくシステムに関する。本発明は、多くのグラフィクス技術と画像処理技術を組み合わせて生産性を向上させる自動化されたユーザフレンドリーなアプリケーションを提供する。

そのアプリケーションは、多くの画像をスキャンする時間を削減し、また、多くの画像の並びを自動的に補正し、さらに、文書のテンプレート中に多数の画像を自動的に配置することによって生産性を向上させることができる。

【0004】本発明は、これらの目的を、

- 1) 複数の独立したオブジェクトを画像中で探索し、
- 2) 認識されたオブジェクト（例えば、矩形）の形状をモデル化し、
- 3) 画像中の各オブジェクトの位置、形状、方向を認識する構造化画像記述を生成することによって達成する。

【0005】画像の分割や構造化画像については、これまで多くの発明と文献で開示されている。それらに関する一部分について以下に要約する。

【0006】1996 年 1 月 16 日に発行されたベナブル（Venable）他の US- A-5,485,568（この文献は、本明細書に組み込まれる）では、複雑なカラーのラスター画像を構造化画像フォーマット、即ち、階層的な装置独立なフォーマットのオブジェクトの集合として表現する方法と装置を開示している。

【0007】“プリプレス・アプリケーションでの画像ワークフローの管理と表現（MANAGING AND REPRESENTING IMAGE WORKFLOW IN PREPRESS APPLICATIONS）”

（（グラフィックアート技術協会（Technical Association of the Graphic Arts（TAGA））、第 1 巻、1995 年、予稿集 pp.373-385、ベナブル他））では、構造化画像を使ってプリプレス・ワークフローを管理することを教唆している。ギャングスキャンニング等の処理は、スキャナー・プラテン上に大雑把に配列された複数の写真を入力する手段として記述されている。

## 【0008】

【課題を解決するための手段】本発明は、入力デジタル画像を処理して複数のオブジェクトの特徴を抽出する方法を提供するもので、入力画像中で少なくとも 2 つのオブジェクトを認識し、各オブジェクトの境界を表現する形状をモデル化し、オブジェクトを特徴づける記述を生成する処理を備え、その記述はさらに画像の他の属性を特徴づける。

【0009】本発明の他の態様は、デジタイズされた入力画像を入力するプログラミング可能なコンピュータを含む画像処理装置を提供することであり、そのコンピュータは、入力画像を格納するフレームバッファメモリと、コンピュータが画像処理を実行するのに適したプログラムコードを格納するプログラムメモリとを備える。その画像処理は、デジタイズされた入力画像中で複数のオブジェクトを識別し、オブジェクトの境界を表現する形状をモデル化し、オブジェクトを特徴づける記述を生成する処理を備える。

【0010】本発明の 1 つの態様は、デジタル画像処理での基本的な問題、即ち、デジタル画像処理で複数のオ

プロジェクトを識別する問題を扱うことである。さらに、この態様は、この問題を緩和する画像処理技術の発見に基づく。ここで記述されている技術によって、ユーザは複数の文書を都合よく 1 回のスキャニング処理でスキャンでき、それらの文書をデジタイズされた画像中で独立なオブジェクトとして認識することによって、自動的に分離することができる。本発明の他の態様は、デジタイズされた画像の構造化画像表現を自動的に生成することによって、画像オブジェクトを独立に簡単に抽出して処理できることである。

【0011】上述の技術は、スキャニング処理効率を向上させるには都合のよいものであり、多数の文書を 1 回でスキャンできる。さらに、その技術によって、ユーザの介入なしにオブジェクトの物理的属性（例えば、位置、形状、方向）を自動的に特徴づけることができる。

【0012】

【発明の実施の形態】本発明の概要の理解のために図を用いる。図全体渡って、同じ参照番号が同じ構成要素を示すために使われる。本発明を説明するために以下の用語を用いる。

【0013】複数のピクセル間に他のピクセルがない場合、その複数のピクセルは画像中で“近傍”または“隣接する”ものであり、その複数のピクセルは近傍の評価基準に適合する。もし、その複数のピクセルが矩形であって 2 次元画像中の行と列にあると、その評価基準に従って、各ピクセルは 4 つの近傍ピクセルまたは 8 つの近傍ピクセルを備える。

【0014】2 つの近傍ピクセルの間にエッジが発生するかどうかを判定する適切な評価基準に照らして、それらの 2 つの近傍ピクセルが十分違うピクセル値をもつと判定された場合、“エッジ”が画像中で発生する。“エッジピクセル”や“境界ピクセル”という用語は、2 つの近傍ピクセルのうちの 1 つかまたはその両方に適用され、その間にエッジが発生する。

【0015】“画像特徴”または“特徴”は、画像で測定可能な属性である。画像を定義するデータを使って特徴を示すデータを作ることによって、特徴を“測定”することができる。もし、特徴が発生する各時点でだいたい同じ結果が得られるような方法でその特徴を測定すると、“画像”に対する特徴が測定されたことになる。

【0016】第 1 の画像の“型（バージョン）”は、第 1 の画像を定義するデータ項目を用いて作られる第 2 の画像である。その第 2 の画像は第 1 の画像に等しいか、または、第 1 の画像を定義するデータを変えたり、第 1 の画像のピクセル値を変えるその他の処理から引き起こされる解像度の損失によって変調される。

【0017】“画像入力装置”は、画像を入力し、画像の型を定義するデータ項目を作る装置である。“スキャナー”は、文書をスキャンする等のスキャニング処理によって画像を入力する装置である。

【0018】“画像出力装置”は、画像を定義するデータ項目を受け取り、その画像を出力として提供したり描画する装置である。“ディスプレイ”は、人間が視覚できる形態の出力画像を提供する画像出力装置である。そして、“プリンタ”は、人間が視覚できるハードコピーの形態の画像を出力する画像出力装置である。

【0019】図 1 と図 2 は、本発明が特定の用途として使われているシステム 20 を示す。システム 20 は、スキャナー 26 のプラテン上に配置された文書（原稿）24 の画像を表現するデジタルデータを入力するコンピュータ 22 を備える。コンピュータ 22 は、まずスキャナー 26 からのデジタル入力データを、画像を連続的にアクセスできるメモリ 52（例えば、RAM や磁気ディスク記憶装置）に格納する。また、デジタルデータに加えて、メモリ 52 は、本発明に係る画像処理工程をプロセッサに実行させる指示を出すオブジェクトコードを格納するプログラムメモリを含む。コンピュータ 22 は、キーボード、キーパッド、マウス、トラックボール、スタイラスまたはそれと等価なポインティングデバイス等の 1 つかそれ以上のユーザ入力装置 30 を含むユーザインターフェイス（U/I）28 と連携する。

【0020】また、システム 20 の一部は、商業ベースで入手可能な多数のプリンタで使われるレーザ駆動のゼログラフィック印刷エンジンを含むプリンタ 34 等の画像出力装置である。好適な実施の形態によれば、システム 20 はスキャナー 26 から入力されたデジタル画像データを処理し、また、プリンタ 34 によって描画され、メモリ 52 に格納され、及び／又はネットワーク 40 を介して他の装置に伝送される出力ファイルを生成するためにプロセッサ 50 で実行される画像処理ソフトウェアを用いる。スキャナーのプラテン上に置かれた文書は複数の写真とサブストレート（基体）表面上でマークによって表現される他のオブジェクトを含むことができ、また、そのオブジェクトを単一のスキャニング処理（操作）でスキャン可能であることが理解される。例えば、特定の実施の形態は、スキャナー 26 のプラテンに接触したりオーバーラップしないように並べられた複数の写真を表現する単一スキャン画像である。本発明によれば、各元の写真の位置、形状、回転角を自動的に認識することが望まれる。

【0021】スキャナーのプラテン上に並べられた複数の異なる写真をスキャンすることによって生成された入力画像が得られると、本発明は少なくとも各写真の位置、形状、方向角を自動的に認識する。図 3 のフローチャートで示されるように、入力画像の処理中にコンピュータ 22 によって実行されるプロセスは、3 つの一般的ステップを備える。まず、ステップ 100 では、画像中のオブジェクトが探索されて、通常、そのオブジェクト境界が識別される。オブジェクトが探索されると、ステップ 200 では、オブジェクト形状をモデル化する。オ

プロジェクトが探索されてその形状がモデル化されると、ステップ300では、その画像とオブジェクトを表現する構造化画像を生成する。その構造化画像は好適には画像データ自体を表現するデータだけでなく、各オブジェクトの位置、形状、方向やそれらの組み合わせを表現するデータを含む。別の方法では、その出力は、画像情報を検索可能な形態で格納するためのページ記述言語フォーマットまたはそれと同等のフォーマットである。

【0022】本発明の好適な実施の形態では、スキャンされた入力画像（または、それより低解像度の画像）がメモリフレームバッファ（RAM）にロードされ、前述のステップに従って分析される。ここでは、説明の便宜上、複数のオブジェクトは互いに隠し合うことがなく、その画像の背景は連続していると仮定する。この単純化された仮定は説明の便宜上のためであって、本発明の範囲を制限するものではない。当業者であれば、本発明をその仮定で必要となる範囲以上に拡張することができる。

【0023】図4のフローチャートで示されるように、オブジェクト位置検出ステップ100では、入力画像102の背景領域をまず識別し、背景領域104を特徴づけ、その背景領域の特徴をシードとして使用し、適応シード・フィル・アルゴリズム106によってその背景領域を表現するピクセル全部を確認する。背景ピクセルは、どのオブジェクトとも関連しないピクセルである。より簡単に言えば、オブジェクトの外側の領域を表現するピクセルであり、そのピクセル値は“背景”によって制御され、その背景に対してオブジェクトがスキャン中に探索される（例えば、プラテンカバーの下面に）。一つの実施の形態では、スキャン画像の左上隅の小領域の色平均を背景色の初期推定値として使用する。別の方法では、ファレル（Farrell）のUS-A-5,282,091“文書の背景レベルを決定するプログラマブル装置”等で開示されているような他のサンプリング処理が行われる。

【0024】ステップ104で背景色が特徴づけられると、好適には適応アルゴリズムが適用され、背景色を制御し、オブジェクトを正確に識別する。本発明に適したシード・フィル・アルゴリズム（種充填アルゴリズム）の一例は、グラフィクスジェム I（Graphics Gems I），A. グラスナー Ed.，アカデミックプレス，pp. 275-277，1990 に記載されている。プラテン領域上の照明変動に起因して背景ピクセルでは大きな色変動が起こるために適応アルゴリズムが必要となる。適応シード・フィル・アルゴリズムは、画像の左上隅等の背景によって特徴づけられる初期シード点を用いてスキャンされたカラー画像データに適用される。一般に、その適応シード・フィル・アルゴリズムでは、背景ピクセルとして識別された連続する全ピクセルを示すマスクを用いて、2値のフレームバッファを塗りつぶす（充填する）。ステップ112に示される簡単な実施の形態では、現在の平均

背景ピクセル値から僅かに離れた距離  $\epsilon$  内にピクセルの色があれば、そのピクセルを背景ピクセルと考える。この距離は、赤、緑、青（RGB）の色空間のユークリッド距離として使って計算される。

【0025】 $d = \text{SQRT}((P_r - \text{AdAvg}_r)^2 + (P_g - \text{AdAvg}_g)^2 + (P_b - \text{AdAvg}_b)^2)$

ここで、 $P_r$ ， $\text{AdAvg}_r$  はそれぞれ、テスト下のピクセルのRGB成分と平均背景値であり、 $d$  は距離である。 $\epsilon$  の値は、一定であり、一つの実施の形態では実験的に決められる。ステップ112では以下のテストが実行される。即ち、もし、 $d < \epsilon$  ならば、ピクセルPは背景ピクセルで、そうでないならばピクセルPは前景ピクセルである。

【0026】ステップ114では、平均背景色は、背景としてカテゴリーされた最後（最新）のN個のピクセルの平均値をとることによって適応的に修正される。好適には、本システムは次式を使い適応的平均値を計算することで処理効率を上げる。

【0027】 $\text{AdAvg}' = (N' \cdot \text{AdAvg} - \text{AdAvg} + \text{LastVal}) / N$

ここで、 $\text{AdAvg}'$  は修正平均、 $\text{AdAvg}$  は前の適応平均、 $\text{LastVal}$  は背景として識別された最新のピクセル値、 $N$  は平均化処理ウィンドーである。明らかに、これは真の移動（実行）平均（running average）ではないが、その移動平均を適切に追従するものであるため、厳密な移動平均演算より計算効率がよい。他の方法では、 $\epsilon$  の値を適応的に変えることができる。例えば、 $\epsilon$  は、背景等として識別された最新の複数のピクセルの標準偏差に基づいて決められる。

【0028】ステップ120では、全背景ピクセルを識別して背景領域を示す2値マスクを作り、形態素フィルタリング（morphological filtering）を行うことで背景マスク中のノイズ性エッジを滑らかにする。具体的には、好適には形態素閉包フィルタは背景マスクに適用され、単一のピクセルノイズを除去し、オブジェクトのエッジを滑らかにする。続いて、ステップ122では、連続した前景領域を探索してオブジェクトを識別する。オブジェクトは、適応シードフィル処理によって生成された背景マスクをスキャンすることによって識別される（ステップ106）。左上のピクセルから始まり、そのマスク中で背景ピクセルとして分類されなかったピクセルをスキャンライン形式で探索することで、前景オブジェクトに関連するピクセルを識別する。そのシード・フィル・アルゴリズムを背景を識別するために使うと、前景オブジェクトは確実に閉包したものになる。

【0029】ステップ124では、オブジェクト境界はそのエッジをトレースすることによって識別される。前景オブジェクト境界は、オブジェクトのエッジをトレースする点の順序付けセットを提供する単純な8連結エッジ走査を用いてトレースされる。そのようなエッジ処理（操作）では輪郭トレース走査処理を用いて、語かまた

は文字ベースの認識システムに似た方法でチェインコードを生成する。8 連結処理は、例えば、R. ボズノビク他 (R. Bozinovic et al.) による“オフライン草書体手書き単語認識”、I E E E・トランザクション・オン・パターン・アナリシス&マシン・インテリジェンス、11 巻、No. 1 (1989 年 1 月) に記載されている。エッジがトレースされると、マスク中のオブジェクトに関連する全ピクセルは背景としてマークが付けられ、次の時点では処理されない。そして、前景オブジェクトリストにそのオブジェクトが加えられ、次に、ステップ 126 で示すように、ステップ 122 でのスキャニング処理を続ける。全オブジェクトを識別する前景スキャニングが完了した後、ステップ 130 では、識別されたオブジェクトを見直す。多くの場合、スキャン画像は不要な前景オブジェクトを含み、そのようなオブジェクトは、このステップでオブジェクトリストから削除される。一つの実施の形態では、オブジェクトリストの見直しにより、画像らしくない小さなオブジェクトを削除する。例えば、各画像が関連する年鑑の頁のスキャンでは、テキストキャプションを画像データとして分類しない。そのようなキャプションは多くの小さな周囲オブジェクトから成るため、トレースされたエッジの周囲長を測定することによって、ある長さ以下の周囲をもつオブジェクトを削除できる。その閾値は実験的に決められる。

【0030】ステップ 100 に関連して述べられているように、オブジェクトが探索されると、次のステップであるステップ 200 では、オブジェクトの形状をモデル化する。説明を簡単にする目的で、以下の記述では矩形オブジェクトを扱うが、他のポリゴン (多角形) や、曲線 (例えば、円や楕円オブジェクト) で表現される部分をもつ形状にまで拡張できることが理解できる。その結果、即ち、ステップ 100 からの出力や結果は、好適にはリンク (連結) リスト中の 1 組のエッジ・トレースであり、これによってスキャン画像中の各オブジェクトの境界ピクセルが識別される。これらのトレースは、各オブジェクトを抽出するために使われるが、その方向はまだ決定されていない。オブジェクト抽出の品質を改善するために、オブジェクト・トレースはモデル形状にフィッティングされる (あてはめられる)。次に、方向情報等がそのフィッティング (あてはめられた) パラメータから抽出される。本発明の実施の形態では、オブジェクト・トレースは矩形モデルにフィッティングされるが、他の形状でも可能である。

【0031】エッジ・トレースを矩形にフィッティングする 1 つの方法は最小 2 乗法であり、これにより矩形にフィッティングする。最小 2 乗法によるフィッティングを行うために、エッジ・トレースはまず 4 組の点に分解される。各組は矩形オブジェクトの 4 辺のうちの 1 辺に対応する。この 4 組の点に分解する処理は、以下で説明

される幾つかの方法でなされる。

【0032】第 1 の方法は、2 つの基本的部分、即ち、(a) エッジ点を、単一の線に関する 1 組のビン (bin) に分類 (カテゴリー化) し、(b) それらのビンで回転された形状の認識処理を行う。図 5 は、第 1 の分解方法を詳細に示す。ステップ 204 では、エッジ・トレースに沿った各点での傾斜を計算する。ステップ 204 では、好適には、近傍のエッジ点の小ウィンドーで線形回帰分析を行うことによって傾斜角の計算を行う。例えば、エッジ点の両側にある 2 点で傾斜が決定される。各点の中心を通る線の角度は、各点を中心とする小ウィンドー内で線形回帰分析を行うことで決定される。各回帰分析では、ウィンドー内の 1 点について 4 つの加算、2 つの減算、2 つの乗算、1 つのアークタンジェント演算が必要であるが、回帰分析アルゴリズムが最適化されるとほとんどの加算演算が不要となる。計算の複雑性を減らす好適な実施の形態では、エッジピクセルのサンプルが傾斜角計算とソーティングのために使われるため、エッジピクセルを分類 (カテゴリー化) するために必要な計算量が減る。

【0033】次に、ステップ 206 では、傾斜カテゴリー、即ち、ビンのリストを構築する。エッジに沿った現在の点 (例えば、図 8 の点 B) とその前の点 (例えば、図 8 の点 A) 間の傾斜角の差の大きさを計算することによって、各エッジ点に対する傾斜が分類される。もし、その差が TOLERANCE 値 (一つの実施の形態では実験的に ±5 度に決められている) より小さいならば、その点が前の点と同じ傾斜カテゴリーに割り当てられるが、そうでない場合には新しい傾斜カテゴリーが作られてそこにその点が割り当てられる。図 8 を参照して、上述の処理によって点 A, B, C が第 1 の傾斜カテゴリーに割り当てられ、点 D, E, F, G, H が第 2 の傾斜カテゴリーに割り当てられ、点 I, J がその他の傾斜カテゴリーに割り当てられる。最後に、もし最後のエッジ点の傾斜カテゴリーの傾斜角が第 1 の傾斜カテゴリーとだいたい同じ傾斜角をもつなら、第 1 と最後の傾斜カテゴリー内の全ての点が共に 1 つのカテゴリーに入る。

【0034】ステップ 206 で傾斜カテゴリーが確立されて、データ構造中に格納されると、ステップ 208 では、それらの傾斜カテゴリーがソートされ、また、各カテゴリーに割り当てられたエッジ点の数に基づいて順序付けされる。矩形オブジェクトでは、ほとんどのエッジ点が含まれる最上位の 4 つの傾斜カテゴリーは、その矩形の 4 つのエッジに沿った点に対応するはずである。ステップ 210 では、その最上位の複数の傾斜カテゴリーが選択される。六角形のオブジェクトでは最上位の 6 つのカテゴリーが使われ、同様に、三角形のオブジェクトでは最上位の 3 つのカテゴリーが使われる等々が理解されるだろう。

【0035】別の方法では、ステップ 208 と 210

は、傾斜角カテゴリー、即ち、そのピンを単純に削除したり統計的に削除することにより処理するステップに置き換えることが可能である。ここで、エントリのほとんどないカテゴリーは削除される。例えば、実験的（経験的）に決められた 5 ピクセルの閾値が使われると、共通の角度で 5 個を超えるピクセルを有するピンが残される。続いて、カテゴリーの平均角度が、特定のカテゴリーに割り当てられた全ての点の単純線形回帰分析によって決定される。決定された平均角度でさらにカテゴリーを改良することによって、実質的に共通の角度をもつ複数のカテゴリーを結合できる。特に、各カテゴリーがチェックされて、もし近傍のカテゴリーが実質的に共線的であれば、それらのカテゴリーは結合される。従って、残りの各ピン、即ち、残りの各カテゴリーはエッジに沿った共線点の組を表わす。残りの各傾斜角カテゴリーに割り当てられたエッジ点は、矩形の 4 辺に分解されたエッジ・トレースを表わす。ここで、このどちらか一方を“フィルタリング”処理、即ち、カテゴリーの整理（詳細化）を行なって、オブジェクトの実エッジを表わすカテゴリーを識別することが理解できる。従って、カテゴリーを整理（詳細化）する同様な方法が検討される。

【0036】オブジェクト境界を特徴づける第 1 の方法は、各エッジ点での平均傾斜の測定に起因して計算的に集約的である。前述の別の実施の形態では、処理速度を改善するためにエッジ・トレースをサンプリングすることで、処理され分類される点の全体数を削減することができる。

【0037】さらに、順序付けされたカテゴリーを分析することで形状を識別できることが理解できる。例え

- a)  $\theta'_0 = \arctan^2(L_1, L_0)$  長軸の半分の角幅  
 $\theta'_1 = 90 - \theta_0$  短軸の半分の角幅  
 $\theta_0 = \theta'_0 + \text{TOL}$  ここで、かどをさけるために  $\text{TOL} = 0.95$   
 $\theta_1 = \theta'_1 + \text{TOL}$  ここで、かどをさけるために  $\text{TOL} = 0.95$
- b) 範囲 1 :  $(\theta + \theta_0)$  から  $(\theta - \theta_0)$   
 範囲 2 :  $((\theta + 90) + \theta_1)$  から  $((\theta + 90) - \theta_1)$   
 範囲 3 :  $((\theta + 180) + \theta_0)$  から  $((\theta + 180) - \theta_0)$   
 範囲 4 :  $((\theta + 270) + \theta_1)$  から  $((\theta + 270) - \theta_1)$

角範囲が決定されると、ステップ 266 では、上述の 4 つの角範囲（マスの中心に関して）の各々にあるエッジ点の全てを検出する。これによって、矩形の各辺に対応するエッジ点を識別する。この技術が、上述の第 1 の方法よりも、エッジノイズに敏感ではないことは理解できる。

【0040】エッジ・トレースが 4 つの点の組に分解されると（各組は矩形の 4 辺の 1 辺に対応する）、ステップ 280 では、矩形にそれらの点をフィッティングするための最小 2 乗演算がなされる。矩形は以下の方程式で

ば、第 3 と第 4 のカテゴリー間で点の数に統計的に大きな差がある場合や第 4 のカテゴリーがない場合は、三角形オブジェクトであることを示す。

【0038】図 6 には、オブジェクト形状をモデル化する第 2 の方法を示す。ステップ 202 でエッジトレースリストデータを検索後、ステップ 252 では、オブジェクトのマスの中心（質量）の中心を計算する。オブジェクトのマスの中心を計算するための多くの既知の方法があるが、矩形オブジェクトの場合、エッジ点の (x, y) 座標の平均をとる方法が簡単な方法である。次に、ステップ 254 では、マスの中心に最も近いエッジ点が探索される。その最も近い点は矩形の長辺の近似中心である。再び図 8 を参照して、マスの中心 (C of M) から中心点 (L, / 2) までの角  $\theta$  は、矩形の近似回転角 ( $\theta$ ) である。

【0039】ステップ 256 で回転角を決定すると、ステップ 258 ではそれを使って矩形の短軸（短辺）の近似の長さを決定する。特に、そのマスの中心から、角範囲  $\theta - \Delta A$  から  $\theta + \Delta A$  にある全エッジ点の平均位置までの距離が決定される。この距離は、矩形の短軸の長さ  $L$  の半分の近似測定（寸法）である。 $\Delta A$  は、経験的に決められた約 5 度のオーダの値である。ステップ 260 では、ほぼ同じ方法で長軸（長辺）（ $L_1$ ）を概算する。そのマスの中心から平均位置  $(\theta + 90) + \Delta A$  までの距離は、矩形の長軸の長さ  $L_1$  の半分の近似寸法である。長軸と短軸の方位角と長さが概算されると、ステップ 264 では、辺に関するこれらのエッジ点のみを含む矩形の各辺に対する角範囲（マスの中心について計算された）を計算する。即ち、

定義される相互に垂直な 4 辺で記述される。即ち、

$$y = \alpha_0 + \beta x$$

$$y = \alpha_1 + \beta x$$

$$y = \alpha_2 + \beta x$$

$$y = \alpha_3 + \beta x$$

ここで、 $\beta_i = -1$  である。最小 2 乗法によるフィッティングにより、以下のあてはめパラメータが得られる。即ち、

【0041】

【数 1】

$$\beta_n = \sum \left( \sum_{i=0}^{n_k} x_{ki} Y_{ki} - \frac{1}{n_k} \sum_{i=0}^{n_k} x_{ki} \sum_{i=0}^{n_k} y_{ki} q \right)$$

$$\beta_d = \left( \sum_{i=0}^{n_1} x_{0i}^2 - \frac{1}{n_0} \left( \sum_{i=0}^{n_1} x_{0i} \right)^2 \right) + \left( \sum_{i=0}^{n_1} y_{1i}^2 - \frac{1}{n_1} \left( \sum_{i=0}^{n_1} y_{1i} \right)^2 \right)$$

$$+ \left( \sum_{i=0}^{n_1} x_{2i}^2 - \frac{1}{n_2} \left( \sum_{i=0}^{n_1} x_{2i} \right)^2 \right) + \left( \sum_{i=0}^{n_1} y_{3i}^2 - \frac{1}{n_3} \left( \sum_{i=0}^{n_1} y_{3i} \right)^2 \right)$$

$$\beta = \beta_n / \beta_d$$

$$\alpha_0 = \frac{1}{n_0} \sum_{i=0}^{n_0} y_{0i} - \frac{\beta}{n_0} \sum_{i=0}^{n_0} x_{0i}$$

$$\beta \alpha_1 = \frac{1}{n_1} \sum_{i=0}^{n_1} x_{1i} - \frac{\beta}{n_1} \sum_{i=0}^{n_1} y_{1i}$$

$$\alpha_2 = \frac{1}{n_2} \sum_{i=0}^{n_2} y_{2i} - \frac{\beta}{n_2} \sum_{i=0}^{n_2} x_{2i}$$

$$\beta \alpha_3 = \frac{1}{n} \sum_{i=0}^{n_3} x_{3i} - \frac{\beta}{n_3} \sum_{i=0}^{n_3} y_{3i}$$

【0042】ここで、 $(x_{ki}, y_{ki})$  は、第  $k$  辺の第  $i$  番目のエッジ点であり、 $n_k$  は第  $k$  辺に関するエッジ点の数である。最小 2 乗法によるフィッティングによって、あてはめパラメータ  $(\beta, \alpha_0, \alpha_1, \alpha_2, \alpha_3)$  が得られると、それらはステップ 282 で矩形のかど

(角) をマークする 4 つの座標ペアに変換される。さらに、矩形オブジェクトの回転角は傾斜パラメータ  $\beta$  によって正確に表現される。

【0043】エッジ・トレースをある形状にフィッティングする他の方法として、高速の画像境界付 (fast image bounding) に対して 2 値モーメントを使う方法がある。ステップ 106 に関して説明されたように、生成された 2 値マスクを使ったり (例えば、適応シード・アルゴリズム)、単純な閾値処理を行うことで背景領域か非背景領域かを示すために各ピクセル値が 0 か 1 である 2

進のビットマップ形式で画像が描画される。2 値マスクを使って 1 つのオブジェクトに対する境界が検出されると、図 7 に描かれた別の実施の形態では、2 次の 2 値モーメントを使い、1 つの形状 (例えば、矩形) をオブジェクトにフィッティングする。

【0044】図 7 には、2 値モーメント境界検出技術の処理工程を示す概略フローチャートが描かれている。ステップ 100 では、オブジェクトのエッジが前述されたように検出され、記録される。これによって、ステップ 290 では、境界のリンク (連結) リスト、即ち、エッジ・トレースと呼ばれるエッジピクセルを入力として提供する。境界リストを使って、2 次モーメントが以下の方程式に基づき効率的に計算される。即ち、

【0045】

【数 2】

$$\mu_k = \sum_i \sum_j p(i, j) i^k j^l = \sum_i \left( \sum_j p(i, j) j^l \right) i^k = \sum_i p_i(i) i^k$$

【0046】ここで、 $p(i, j)$  は、画像座標  $(i, j)$  での画像ピクセル値であり、 $p_i(i)$  は第  $i$  番目のスキャンラインの第  $i$  次モーメントである。オブジェクト境界ピクセルが前に決定されているので処理は簡単になり、また、特定のスキャンラインに対する最右境界ピク

セルと最左境界ピクセルが、1 次 (絶対) のモーメント演算のために使われる。

【0047】続いて、2 次 (中心) のモーメント

$(m_{00}, m_{01}, m_{10}, m_{11}, m_{20}$  及び  $m_{02})$  が 1 次のモーメントと以下の方程式を用いて計算される。即ち、

【0048】

【数3】

$$x_c = -\mu_{10} / \mu_{00},$$

$$y_c = -\mu_{01} / \mu_{00}, \text{ 及び}$$

$$m_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} x_c^{p-r} y_c^{q-s} \mu_{rs}.$$

【0049】2次のモーメントが決定したので、ステップ294では、それらを用いて楕円を特徴づけ、その楕円からオブジェクト回りの境界ボックスを求める。特に、楕円の中心(x、y)、各軸の長さ(aとb)、回転角(θ)が決定される。矩形オブジェクトの境界ボックスは、角θ分回転した、2aと2bの長さの辺をもつ(x、y)を中心とする矩形として決定される。この境界ボックスは、安全な演算マージンをとるため及びオブジェクトの一部のクロッピングを回避するために、オブ

10

ジェクトよりわずかに大きく取られる。もし、より狭い(よりタイトな)境界ボックスが必要であれば、矩形は2αaと2αbの長さの辺で特徴づけられる。ここで、αは√3/2に設定されるか、それよりわずかに小さい値に設定され、それによりエッジをトリミングやクロッピングする(例えば、1つかそれ以上のピクセルのオーダーで)。

【0050】各オブジェクトがある形状(例えば、矩形)にモデル化された後で、例えば、ベナブル他(Venable et al.)のUS-A-5,485,568で開示されているような構造化画像が生成される。その構造化画像は、上述の方法の1つを用いて検出された各オブジェクトに対して1つの“子”構造化画像をもつ。その構造化画像の定義は、スキャン画像のどの矩形がオブジェクトデータを含むかを示す属性と、どんな方位スキューでも補正を行なうために必要な回転角を含む。図9は、前述したプロセスによって生成された構造化画像の一例を示す。ここで、その構造化画像は一对の矩形画像オブジェクトを含む。

【0051】図10に描かれた本発明の一つの実施の形態では、描画されると、全てのオブジェクトは回転が解

20

30

40

50

きる手段が提供される。

【0052】例えば、図1と図10を参照して、ユーザは多数の写真をスキャナのプラテンに配置できる。これらが配置されると、次に、ユーザは図10の領域410から処理を選択して、コンピュータシステム22にスキャナ26によるスキャン処理を開始させる。図10に示されているように、“一括編集 (Gang & Edit)” 412が選択されると、システム20はプラテン24上に置かれたオブジェクトをスキャンし、ユーザインターフェイス画面の領域420に反映されたファイルの詳細を使ってそのファイルにデータを一時的に格納する。例えば、様々な画像オブジェクト(A、B、C、D)が、図10に描かれているような画像中で検出される。その画像がスキャンされると、それは上述したように分析されて、画像オブジェクトが識別される。次に、画像オブジェクトは、自動的に画像を方向付け、位置するために、スマートスキャニングシステムによって操作される。例えば、ユーザインターフェイスの領域430中に描かれた表現等のように画像オブジェクトは所定のテンプレートに自動的に配置されて描画される。テンプレートに関する別の編集能力、例えば、オブジェクトにキャプションを追加する能力、即ち、図示されているようにタイトル432とサブタイトル434を含ませる能力がユーザに提供されることは高く評価される。領域440中のユーザインターフェイスオプションを介して、テキストベースでの編集のための入力となされる。

【0053】また、スマートスキャニングシステムは、ユーザインターフェイスの領域450中に描かれている画像編集能力を備える。画像中の各オブジェクトが識別されると、複数のオブジェクトを分離し、それを用いて複数の別々の画像を生成し、その画像を個々に処理することができる。従って、領域430のテンプレート中に配置された個々の画像オブジェクトは、個別に選ばれるか、操作されるか、スケーリング(ボタン452で)されるか、回転(ボタン454で)されるか、クロッピングされる(ボタン456で)。前述のオブジェクト識別方法の結果として本システムが自動的に行なう処理に対して、スケーリング(拡大縮小)、回転、クロッピングの各処理をさらに加えたことは高く評価される。

【0054】例えば、画像オブジェクトをスケーリングするために選択を示す線影で描かれた画像スキャンボタンによって、ユーザはカーソルを移動させて1つのオブジェクト(例えば、画像オブジェクトD)を選択し、そのオブジェクトの辺、または、そのかどをドラッグすることができる。オブジェクトの編集を容易にするために、例えば、画像オブジェクトDの境界に描かれている制御点(436)がユーザインターフェイスの設計者には公知の方法で使用されることができる。

【0055】所定のテンプレートを使って、画像オブジェクトを、文書やそのページの相対的位置に自動的に配



置できることは注目に値する。このようなテンプレートは構造化画像定義形式であってもよく、そのテンプレートを使って、生成される構造化画像に対して異なるレイアウトを施すことができる。従って、“デジタル写真アルバム”に写真を配置したい家族は、ユーザインターフェイスの領域 4 3 0 に示されたページに似たページを記述するテンプレートを作ることができる。そして、そのテンプレートを使って、より大きい文書画像中に個々の画像、即ち、複数のオブジェクトを自動的に構成することができる。

【0 0 5 6】好適な実施の形態では、出力は、ベナブル他(Venable et al.)のUS-A-5,485,568で開示されている構造化画像出力フォーマットである。構造化画像の1つの重要な特徴は、画像処理の複数の操作をそれらの記述内に記録することができることである。これは、構造化画像が単純なオブジェクトのスキュー除去属性以外に複数の画像処理操作を含むことができることを意味する。例えば、識別される個々のオブジェクトを改善(強調)するために、複数の自動画像改善(強調)操作を構造化画像中に含ませることができる。

【0 0 5 7】ユーザが望む条件で、ウィンドー 4 3 0 中に“ページ”が構成されると、ユーザは“編集画像のセーブ(Save Edited Image)”ボタン 4 6 0 を選択することによって、その画像をセーブさせることができる。そしてもっと重要な点は、ユーザが、組み立てられた(合成)ページ(単数又は複数)を印刷するか、または伝送することができることである。

【0 0 5 8】本発明に係るさまざまな実施の形態が、このスマートスキャニングシステムに関連して説明されたが、画像の獲得と構成されたページの印刷または伝送をネットワークを介してか、または、ウォークアップ型(walk-up)デジタルコピー機で達成することができる。例えば、ユーザは写真をフィルムプロセッサによって自動的にスキャンさせ、ネットワーク経由でユーザにデジタルサイズされたスタンプシート(stamp sheet)が送られる。構造化画像フォーマットであるスタンプシートは、スマートスキャニングシステムを使って処理され、1つ

かそれ以上のオブジェクトを各ページにもつデジタル写真アルバムの複数のページを作ることができる。

【0 0 5 9】要約すると、本発明は、デジタル入力画像を処理して複数のオブジェクトを特徴づける方法とその装置である。本技術は、背景ピクセルと前背景ピクセルを特徴づけることによって入力画像中の少なくとも1つのオブジェクトを識別し、2つの一般的方法のうちの1つを使ってオブジェクト境界を表現する形状をモデル化し、オブジェクトを特徴づける記述を生成する技術を含む。その記述は、オブジェクトの形状と位置だけでなくオブジェクトの回転やスキュー情報も含む。

【図面の簡単な説明】

【図 1】本発明の一つの実施の形態の画像処理システムを構成する装置の図である。

【図 2】図 1 のシステムを備えるさまざまな構成要素のブロック図である。

【図 3】本発明に係る図 1 と図 2 のシステムを実行させる概略処理工程を示すフローチャートである。

【図 4】本発明に係るさまざまな実施の形態によって実行される処理工程を示す詳細なフローチャートである。

【図 5】本発明に係るさまざまな実施の形態によって実行される処理工程を示す詳細なフローチャートである。

【図 6】本発明に係るさまざまな実施の形態によって実行される処理工程を示す詳細なフローチャートである。

【図 7】本発明に係るさまざまな実施の形態によって実行される処理工程を示す詳細なフローチャートである。

【図 8】デジタル文書の一部を一例を示す図である。

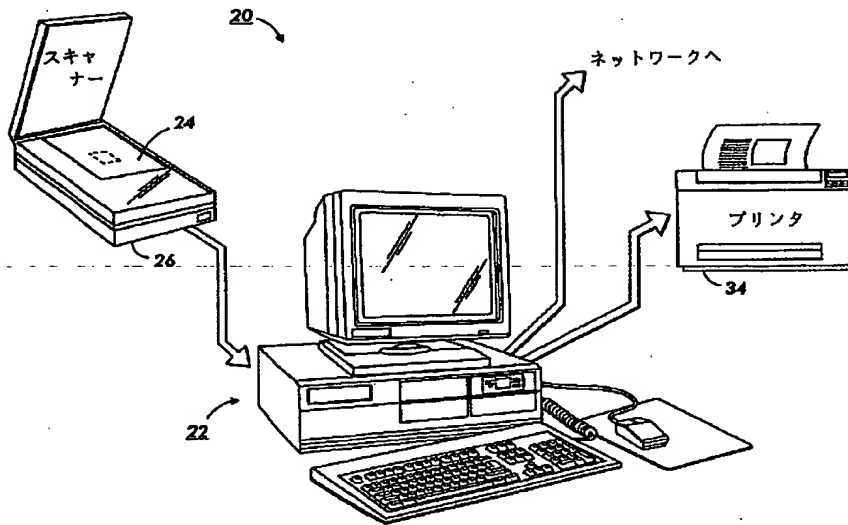
【図 9】入力画像が本発明によって処理されるとき図 1 のシステムの出力を示す図である。

【図 1 0】本発明の一つの実施の形態に関連するユーザインターフェイス画面の一例を示す図である。

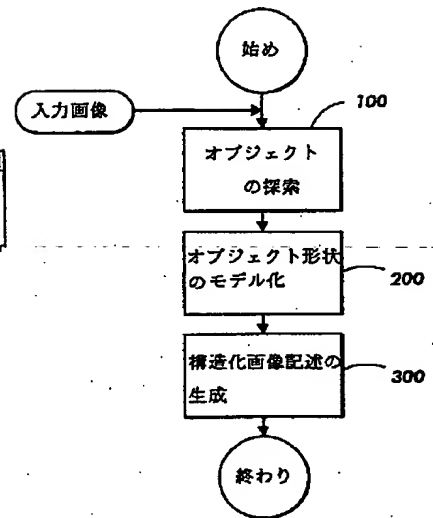
【符号の説明】

2 0 システム  
2 2 コンピュータ  
2 6 スキャナ  
3 4 プリンタ

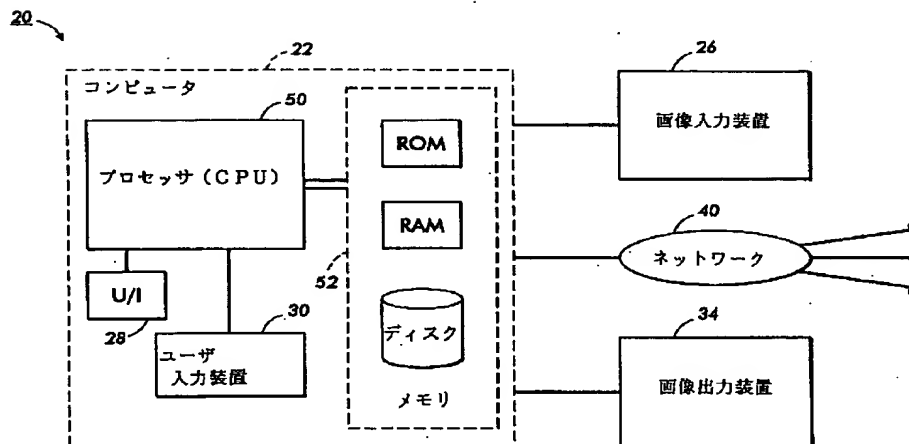
【図 1】



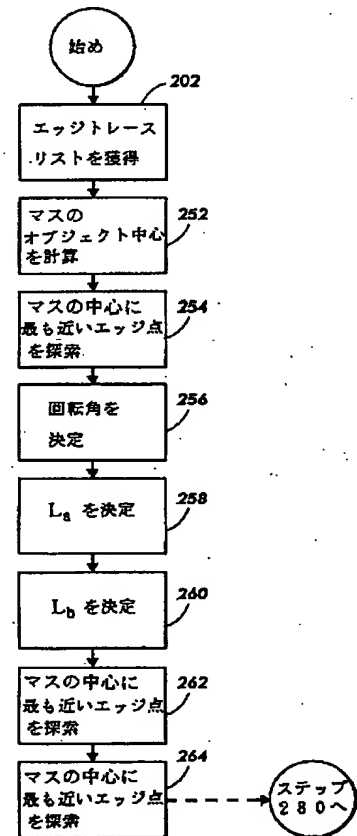
【図 3】



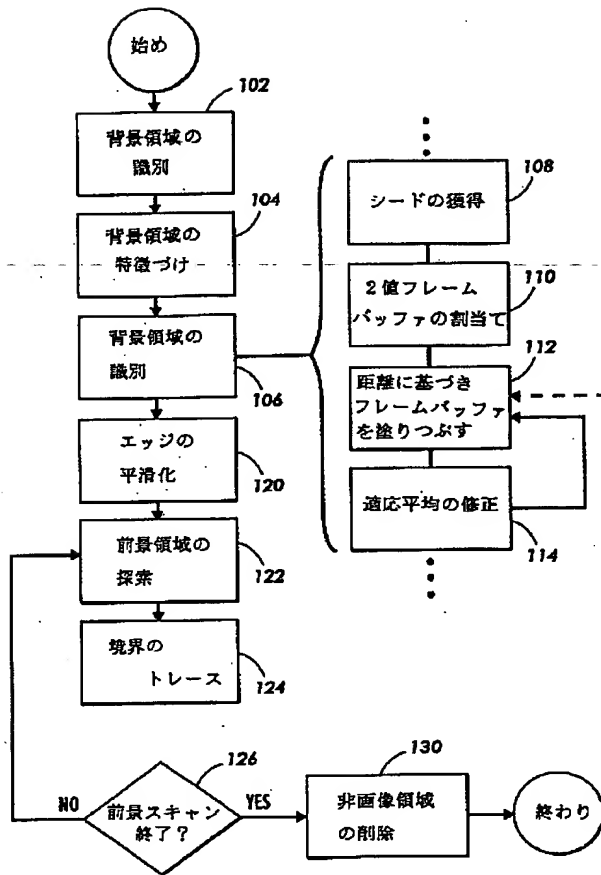
【図 2】



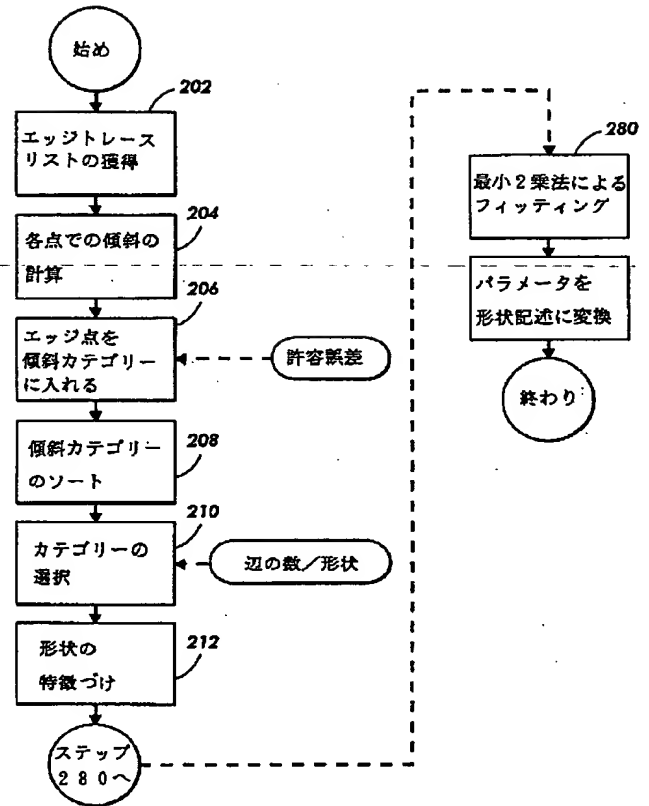
【図 6】



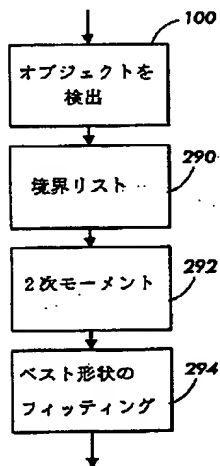
【図4】



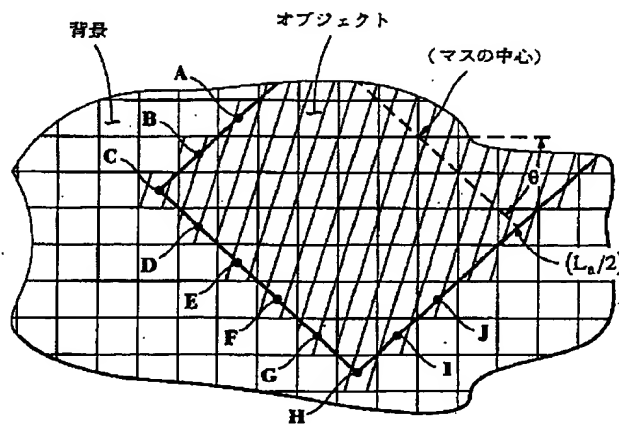
【図5】



【図7】



【図8】



【図9】

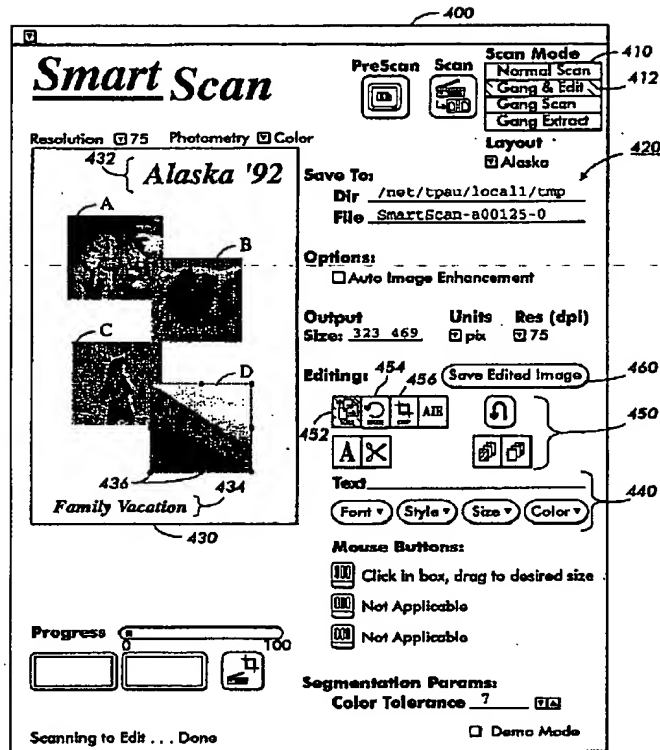
```

# Structured Image ver1.0
sid : smp = {
  aspect_ratio = 1.0 ;
  representation = {
    format = ipd ;
    data = [
      merge = {
        xy = 0.000000 0.132377 ;
        path = {
          object = (                                     #first detected object
            size = 0.500000 0.735247 ;
            sid : image1 = {
              aspect_ratio = 1.470494 ;
              representation = {
                format = raster ;
                data = "smp.int" ;           # scanned image
                attribute = {
                  selection = $sell ;       # object rectangle
                  derotate = 14.445258 ;    # derotation angle
                } ;
              } ;
            } ;
          } ;
        } ;
      } ;
    ] ;
  } ;
  merge = {
    xy = 0.500000 0.180773 ;
    path = {
      object = (                                     #second detected object
        size = 0.500000 0.638454 ;
        sid : image0 = {
          aspect_ratio = 1.276908 ;
          representation = {
            format = raster ;
            data = "smp.int" ;           # scanned image
            attribute = {
              selection = $sel0 ;          # object rectangle
              derotate = 12.238364 ;      # derotation angle
            } ;
          } ;
        } ;
      } ;
    } ;
  } ;
} ;

selection : sell = Selection: In           # rectangle of second object
rect: Include
0.858490 0.958967

```

【図 10】



フロントページの続き

(72)発明者 リカード エル. デ クエイロツ  
 アメリカ合衆国 14450 ニューヨーク州  
 フェアポート シェニン ラン 32

(72)発明者 ラマナ エル. ラオ  
 アメリカ合衆国 87544 ニューメキシコ  
 州 ロス アラモス トゥエンティ サード  
 ストリート 2003ディー